Comparative methods for distinguishing flakes from geofacts: a case study from the Wenas Creek Mammoth site

Patrick M. Lubinski [a, b, *], Karisa Terry [a], Patrick T. McCutcheon [a, b]

[a] Department of Anthropology, Central Washington University, 400 East University Way, Ellensburg, WA 98926-7544, USA
[b] Resource Management Program, Central Washington University, 400 East University Way, Ellensburg, WA 98926-7544, USA

ARTICLE INFO

ABSTRACT

Archaeologists have long struggled with distinguishing lithic materials modified by humans (artifacts) from natural objects (e.g., geofacts or zoofacts). This problem is especially difficult for finds of small numbers of flake-like lithic specimens, and particularly for very old finds. We attempt to address the artifact versus geofact problem at a paleontology site by employing three systematic and objective tests on the two recovered possible artifacts. First, they are compared with debitage attributes typically expected of artifacts and geofacts based on published experimental and actualistic data. Second, they are compared in terms of nine of these attributes with a toolstone sample from the site excavation matrix. Third, the two possible artifacts are scored for these nine attributes and graphed against the toolstone matrix sample and two samples of flintknapped debitage assemblages. In all three comparisons, the two specimens are more like artifacts than geofacts. While this does not prove the specimens are artifacts, it at least shows they cannot be easily dismissed as the sort of geofacts typically expected in the site matrix. We argue that this distinction is an important first step in the evaluation of possible lithic artifacts.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Fragments of stone resembling chipped stone debitage are found in many contexts, and archaeologists must make a decision on whether they are likely to have been created by human action or other causes. This is a routine problem encountered during pedestrian survey in areas with abundant, naturally-occurring, knappable lithic material (toolstone) on the surface, but also occurs in many excavations. Although the problem is significantly reduced for many researchers when there are large numbers of obvious artifacts accompanying the more uncertain stone fragments, or when there are large samples of the uncertain specimens that might be compared statistically with known cultural (anthropogenic) debitage, the problem remains for those unfortunate enough to find themselves with only small amounts of possible flaking debris. While the small numbers do not by themselves justify dismissal of such sites, many researchers find it is difficult to arrive at definitive conclusions about whether humans made such small samples of simple specimens that could

conceivably have been created by non-human means. So, we are forced either to ignore such finds or else to work towards some means to assess them objectively.

This issue is especially important for very old sites, those similar to or older than the currently-accepted age for the beginning of the archaeological record in a given study area. Worldwide, there has been considerable debate about the earliest chipped stone sites and what evidence is needed to accept them (e.g., Ackerman, 1989; Collins, 1997; Gillespie et al., 2004; Haynes, 1973; Meltzer et al., 1994; O'Connell and Allen, 2004; Roebroeks and van Kolfschoten, 1994; Shea, 2010). In North America for example, several late Pleistocene paleontological sites have yielded small amounts of possible flaking debris (e.g., Byers, 2005; Dunbar, 2006; Fosha et al., 2012; Fuld et al., 2014; Hamilton, 1996; Joyce, 2013; Lubinski et al., 2007; Richards et al., 1987; Wyckoff et al., 2003). Acceptance of such purportedly early North American sites is typically considered to require not only compelling evidence that the specimens were created by humans, but also reliable age estimates and strong associations between these age estimates and artifacts (Dincauze, 1984; Haynes, 1969; Meltzer, 2004). Additional criteria for accepting purportedly early sites elsewhere in the world include strongly associated faunal remains, butchered bone, fire features or burnt artifacts, and hominin remains (e.g., Ackerman, 1989; Roebroeks and van Kolfschoten, 1994; Shea, 2010).

* Corresponding author. Department of Anthropology, Central Washington University, 400 East University Way, Ellensburg, WA 98926-7544, USA. Tel.: +1 509 963 3601; fax: +1 509 963 3215.
E-mail address: Lubinski@cwu.edu (P.M. Lubinski).

In this paper, we focus strictly on evaluation of possible lithic debitage. Although a thorough evaluation of any purported site would require investigation of multiple, independent lines of inquiry, there are a number of finds with little more than a few purported lithic artifacts, and determining if specimens are of human manufacture is a logical first step at any site. Here, we develop a protocol to attempt systematic evaluation of a few fragments of possible chipped-stone debitage from a paleontological site. A similar approach could be taken at a site of any age, but the study site here is a late Pleistocene paleontological find in North America. The goal is to determine if the specimens are more like artifacts or more like *geofacts* (Haynes, 1973) that occur naturally in the site matrix. This is accomplished by comparing the possible artifacts to (1) expectations for cultural materials and geofacts garnered from previous experimental and actualistic studies, (2) the toolstone present in the site matrix, and (3) two flintknapped debitage assemblages. It is understood that such an exercise does not test all plausible hypotheses for the origin of the specimens, and that it may provide a probabilistic rather than definitive answer to the question of human manufacture, but we view this as a helpful initial step in evaluation.

1.1. Distinguishing geofacts from lithic debitage

For more than a century, archaeologists have struggled with distinguishing lithic materials modified by humans from natural objects (e.g., Andrefsky, 2013; Ascher and Ascher, 1965; Barnes, 1939; Buehler, 2003; Duvall and Venner, 1979; Ellen and Muthana, 2013; Gillespie et al., 2004; Grayson, 1986; Munson and Frye, 1965; Payen, 1982; Peacock, 1991; Reeves, 1980; Shea, 2010; Warren, 1905). This distinction has been especially controversial for purported early (pre-Clovis) sites in the Americas (Dincauze, 1984; Haynes, 1973; Leakey et al., 1968; Meltzer et al., 1994), just as it was for early "eolith" sites in late nineteenth-century Europe (Grayson, 1986). Even limited to possible flaking debitage alone, a wide variety of attributes has been suggested as either diagnostic or suggestive of human origin since at least the late 1800s (e.g., de Mortillet, 1883:82). This study of features on stone specimens to identify signatures of anthropogenic and natural processes can be viewed as part of chipped stone taphonomy, analogous to taphonomy in faunal studies (see Eren et al., 2011).

Some of the proposed criteria for distinguishing geofacts from artifacts for lithic specimens resembling debitage have been subjected to evaluation by experimental and actualistic studies. Table 1 provides all of the criteria that we could find published that included explicit testing data, namely experimental flintknapping (Patterson, 1983), experiments that attempt to mimic "natural" flake production (Luedtke, 1986; Nash, 1993), and/or comparisons of known artifact and geofact assemblages (Barnes, 1939; Bradbury, 2001; Peacock, 1991). The table includes both criteria that appear well supported, and criteria that have not been substantiated by these experimental and actualistic studies. Naturally, these studies are few and not fully comprehensive or exhaustive, but they provide some objective means for evaluating the proposed criteria.

The 13 attributes listed in Table 1 focus on features resulting from initial flake removal, or what Andrefsky (2013:418) calls flake blank formation. By this we do not mean initial reduction (e.g., decortication flaking) but rather attributes of unmodified debitage, excluding those from edge modification subsequent to initial flake detachment. This approach was taken because (1) there is much less written on unmodified debitage than on edge-modified specimens in the artifact–geofact literature, (2) this is most relevant for our two possible debitage specimens, (3) this can be addressed with low magnification, and (4) edge modification on debitage is known to occur as a result of post-depositional alteration (Hosfield

Table 1
Tested lithic debitage attributes and cultural interpretation.

Attribute	Typical of geofacts	Typical of artifacts	Reference[a]
Characteristics supported by testing			
A) Identifiable dorsal & ventral surface	Absent	Present	Yes: 2
B) Platform cortex	Present	Absent	Yes: 5 No: 6
C) Differential weathering of flake scars	Present	Absent	Yes: 3, 5, 6
D) Bulb of percussion	Absent	Present	Yes: 5, 6 No: 3, 4
E) Bulb of percussion shape	Diffuse	Pronounced	Yes: 3, 5, 6 No: 1, 4
F) Eraillure scars (Bulbar scars)	Absent	Present	Yes: 4, 5, 6 No: 3
G) Fissures	Absent	Present	Yes: 6
H) Dorsal flake scar count	None or Few	Multiple	Yes: 3, 4, 5, 6
I) Dorsal flake scar orientation	Not parallel	Parallel to medial axis	Yes: 1, 3, 4, 5, 6
J) Dorsal cortex	Present	Absent	Yes: 3, 5, 6
Characteristics not supported by testing			
K) Exterior platform angle	>90°	<90°	Yes: 1 No: 3, 4, 5
L) Platform faceting	Absent	Present	No: 6
M) Ripple lines	Absent	Present	No: 3, 4, 6

Note: All of these attributes are purportedly more common in one class (e.g., geofacts) than the other; thus they are suggestive rather than diagnostic.
[a] References with data suggesting this criterion is helpful (Yes) or not helpful (No): 1 = Barnes 1939; 2 = Bradbury 2001, 3 = Luedtke 1986; 4 = Nash 1993; 5 = Patterson 1983; 6 = Peacock 1991.

and Chambers, 2003; Luedtke, 1986; Patterson, 1983). Although it is beyond the scope of this study, we note that there are some patterns of edge modification that purportedly indicate artifacts. For example, debitage artifacts are expected to have unifacial, patterned, and evenly-sized marginal microflakes, while geofacts are expected to have bifacial, random, and unevenly-sized microflakes (Luedtke, 1986; Patterson, 1983).

The attributes listed in Table 1 have varying levels of support by experimental and actualistic studies. Some have strong and unanimous support by three or more studies, such as (C) differential weathering of flake scars, (H) dorsal flake scar count, and (I) dorsal flake scar orientation. Other attributes appear soundly rejected, such as (K) exterior platform angle and (M) ripple lines. Many attributes had less definitive testing results. We listed all those attributes with a majority of studies showing they were not helpful as "not supported" and the remainder as "supported." Supported attributes include (A) identifiable dorsal and ventral surface, also called single interior surface (see Sullivan and Rosen, 1985), (B), cortex on the striking platform, (C) differential weathering of flake scars, (D) identifiable bulb of percussion/bulb of force, (E) a pronounced as opposed to a diffuse or flatter shape of the bulb of percussion (see Peacock, 1991), (F) an eraillure scar (a small flake scar on the bulb of percussion), (G) fissures, also known as radial striations, radial lines or hackles, (H) multiple flake scars on the dorsal surface (excluding any retouched edge flakes), (I) dorsal flake scars parallel to the medial axis of the flake, and (J) cortex on the dorsal surface.

Note that some attributes may be rare in assemblages for which they are noted as "typical" in Table 1. For example, pronounced bulbs of percussion are listed as typical of artifacts rather than geofacts, even though they may be rare in some cultural assemblages, such as those produced with soft hammer percussion (Cotterell and Kamminga, 1987). The point here is that when they do occur, they are more typical on artifacts than geofacts, based on the studies cited.

Even well-supported criteria in Table 1 are suggestive rather than diagnostic of human or natural origin, meaning that they can occur in both, but are more commonly found in one than the other. Thus, the strongest cases for cultural origin would seem to be specimens with large numbers of these culturally-suggestive attributes, and weak cases would be those with few culturally-suggestive attributes. Some authors (Peacock, 1991; Staley, 2006) have built upon this idea by using a scoring system with points awarded for each attribute suggestive of artifacts, and sometimes points detracted for attributes suggestive of geofacts. Both of these authors then graphically compare the distribution of scores of unknown samples to scores of samples of known origin (i.e. known artifacts and geofacts).

Another category of specimen attributes not listed in Table 1 is metrics, such as measurements of length, width, and thickness of complete flakes. It has been suggested that humanly-produced flakes are relatively longer, narrower, and thinner than geofact flakes (e.g., Staley, 2006). However, several studies show a wide range of overlapping length/thickness values (Bradbury, 2001; Luedtke, 1986) that preclude separation of individual flakes but might allow separation of assemblages. One could argue that metrics and all other specimen-based attributes should be compared between assemblages rather than between individual specimens. This would allow for statistical comparison of samples of unknown origin to those of known cultural and natural origin (e.g., Duvall and Venner, 1979). While this is undoubtedly helpful where large samples are present, there are a number of sites, including the study site, where the possible lithic artifacts are few in number. The fact that there are few does not in itself prove they are not cultural, and consideration of individual specimen attributes is necessary.

Any attributes selected for distinguishing artifacts from geofacts need to be tempered with knowledge of geological context. As noted by Nash (1993: 135) "criteria that are appropriate for distinguishing artifacts from naturefacts at one context may not be appropriate for another." For example, flake attributes commonly found in geofacts from a rock fall experiment (Nash, 1993) might provide good criteria for a rockshelter site with an ébolois site matrix, but are less appropriate for a dissolution rockshelter site filled with eolian sand. The strongest use of lithic specimen attributes would employ criteria found useful for distinguishing geofacts from artifacts in similar depositional environments, with similar raw materials. Unfortunately, such data do not exist for many depositional environments and raw materials, including cherts in fine-grained colluvium like the study site.

The geological context of a find clearly is important in other ways as well. As an obvious example, a single, cortex-covered flake discovered in a streambed composed of the same raw material is much less compelling than an identical object found in overbank clay sediment far from any other specimen of the same material. Pertinent aspects of context include the lithology of the specimen compared to the site matrix and depositional environment of the find and its matrix. The lithology concern has long been considered, with local raw materials seen as more consistent with geofacts, and exotic raw materials with artifacts (Patterson, 1983; Peacock, 1991). As with the attributes in Table 1, the lithology criterion is only suggestive, as there certainly are artifacts fashioned out of local materials (e.g., quarry sites), and geofacts of apparently exotic materials (e.g., glacial erratics).

The depositional environment concern is mostly about the energy implied from the sediments in which finds are embedded. Specimens are often considered geofacts if found in dynamic, high-energy depositional environments that might have produced them naturally, such as talus cones at the base of cliffs (Meltzer et al., 1994), in glacial till (Oakley, 1967:12; Wilson and Burns, 1999), or more generally in coarse sediments like river gravels (Roebroeks and van Kolfschoten, 1994). However, like all the other criteria discussed here, depositional environment is only suggestive, as legitimate artifacts certainly are known from high-energy deposits (e.g., handaxes found in French river gravels).

Another context to consider is behavioral/technological organization, that is, the expected nature of the lithic assemblage at the site. If the lithic artifacts at a find are consistent with the expectations for the site type, then the archaeological nature of the find would be strengthened. For a find like the Wenas Creek Mammoth site, the expected activities would presumably relate to killing, butchering, or scavenging of the mammoth and/or bison remains if the site is indeed archaeological rather than purely paleontological. Based on reported megafauna kill/butchery sites (e.g., Frison and Todd, 1986; Sanchez et al., 2014; Saunders and Daeschler, 1994; Sellards, 1938), one might expect hafted bifaces (projectile points/knives), unifacial knives or scrapers, and/or debitage from tool resharpening or damage. Unless the site also includes an associated camp, one would not expect large numbers of artifacts; for example the 40 m^2 bonebed excavation at El Fin del Mundo yielded 4 points and 21 retouch/resharpening flakes in situ or in 1/8 inch (3.2 mm) screen (Holliday, personal communication; Sanchez et al., 2014). Of course, there could be unexpected artifact types such as unmodified early stage debitage that was simply lost or discarded here, but these provide a much more ambiguous connection that is harder to understand in the context of kill, butchery, or scavenging. The latter is the case for the study site used here, where there are two specimens of possible human manufacture without a strong connection to expectations of human activity at the locale.

1.2. The study site

The site for which we develop the protocol is the Wenas Creek Mammoth site (45YA1083), discovered during 2005 construction of a private road in central Washington State (Fig. 1). It was the subject of annual Central Washington University summer investigations from 2005 to 2010, under direction of the senior author. The excavations recovered remains of one mammoth (*Mammuthus* sp.) and one bison (*Bison* sp.) from a single stratum. Two lithic specimens (catalog no. 176 and 327) found in the excavations are possible artifacts (Lubinski et al., 2009). These two possible artifacts are evaluated in our analysis.

The site lies on a laterally-discontinuous bench on an interfluvial ridge between Wenas Creek and the Naches River. The northeast-facing bench is approximately 170 m below the top of the interfluve and 21 m above the Wenas Creek floodplain (Lubinski et al., 2007). As such, it is ~95 m above and roughly 10 km northwest of the furthest reaches of the Missoula outburst flood deposits as mapped by Waitt (1980) and extended by Lillquist et al. (2005). The site lies in a locality mapped as a Quaternary landslide deposit modifying a ridge composed of Ellensburg Formation (Bentley and Campbell, 1983) sediments of Miocene age (Smith, 1988). The Ellensburg Formation near the site is primarily fluvial and lahar-derived fines and gravels dominated by pumiceous dacite, andesite, and basalt (Bentley and Campbell, 1983), but this formation is also known to contain silicified materials such as petrified wood and cherts in the general Columbia Basin region (Miller and Powell, 1997). These Ellensburg Formation silicified materials provide knappable toolstone in pieces up to 26 × 20 × 17 cm in size exposed in ephemeral drainages well uphill of the site.

The bones and possible artifacts derive from a ~20–50 cm thick gravelly silt loam diamicton (Stratum II; Fig. 2), interpreted as colluvium. This stratum is quite fine grained, composed of 10% gravel and 90% mud based on sieve analysis from a column at the south edge of the excavation. Toolstone in the excavated matrix
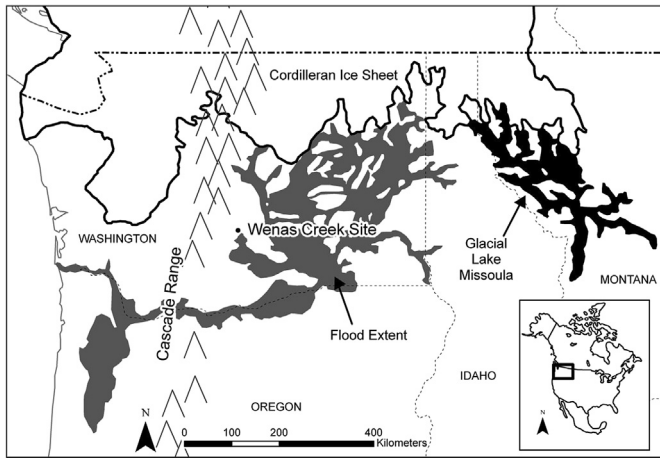
Fig. 1. Site location compared to late Pleistocene landscape features, modified from Waitt (1985:Fig. 1). The gray area labeled "Flood Extent" refers to the maximum area swept by Glacial Lake Missoula outburst floods and/or inundated by Glacial Lake Columbia.

occurs at a low density of 46.9 pieces/m^3, based on a toolstone sample (see Section 3.2 below). Stratum II is overlain by a silty loam ~60−80 cm thick (Stratum I) interpreted as loess, and underlain by more than 180 cm of bedded sands and gravels interpreted as sidestream alluvium (Stratum III).

The two possible artifacts were found under slightly different situations during the excavations (Fig. 3). Catalog no. 176 (FS 261) was recovered in situ in 2006 from about mid-stratum II and about 15 cm above the nearest bone (Fig. 2), an undated mammoth-size metapodial. Catalog no. 327 (FS 479) was recovered in the screen in 2007 from the base of Stratum II near its contact with the underlying, higher energy, alluvial Stratum III. The possible artifacts are illustrated and described in detail in Section 3.1 below.

Ages of these strata and their constituents have been reported in detail by Lubinski et al. (2014). While age estimates and association of age estimates with possible artifacts are not strictly relevant to the independent evaluation of the possible artifacts in this paper, we summarize the results here for the interested reader. There are

currently eight collagen AMS radiocarbon dates of bone recovered in Stratum II, and 14 infrared-stimulated luminescence (IRSL) dates on sediments (three on Stratum I, seven on Stratum II, and four on Stratum III). The eight radiocarbon dates provide a mean pooled age of 13,874 ± 24 RCYBP and a pooled 2σ range from 16,782 to 17,128 CALYBP using the CALIB 6.1.1. radiocarbon calibration program (Stuiver et al., 2012). The seven IRSL dates from Stratum II provide individual 2σ ranges from 27.0 to 6.1 ka, five of which overlap with the radiocarbon dates. These dates indicate the mammoth and bison at the site died about 17 ka and were buried in sediment, most of which was deposited at about the same time, but some of which was deposited later. These two depositional age components are indicated in 94 single-grain IRSL age estimates pooled from the four samples collected near catalog no. 176; the two resolved components are 16.8 ± 0.9 ka (80% of grains) and 5.1 ± 0.5 ka (20% of grains).

The likely source for Stratum II is mass wasting of Ellensburg Formation deposits from uphill of the site. Thus, angular rocks resembling flakes in Stratum II could conceivably be derived from unaltered Ellensburg Formation gravels deposited by streams and lahars more than 5 million years ago. Alternately, they could be derived from collision of Ellensburg Formation gravels as they traveled downslope by mass wasting and were deposited onsite about 17 ka. Subsequent to that initial deposition, additional rocks may have been moved into Stratum II through graviturbation or cryoturbation, although there is no direct field evidence of these processes. Angular rocks resembling flakes may also have been introduced by bioturbation, based on the few krotovina visible in unit profile walls in Stratum II, although a surface survey extending more than 50 m beyond the site failed to locate any chipped stone artifacts or indeed any toolstone pieces more than 5 cm in size. Nonetheless, there clearly is potential for Holocene intrusive material at the site, especially given the 5.1 ± 0.5 ka IRSL sediment component (Lubinski et al., 2014).

2. Methods

The two possible artifacts were evaluated by comparing them against (1) theoretically expected cultural debitage and geofact features as described in the Section 1.1, (2) attributes of a systematic
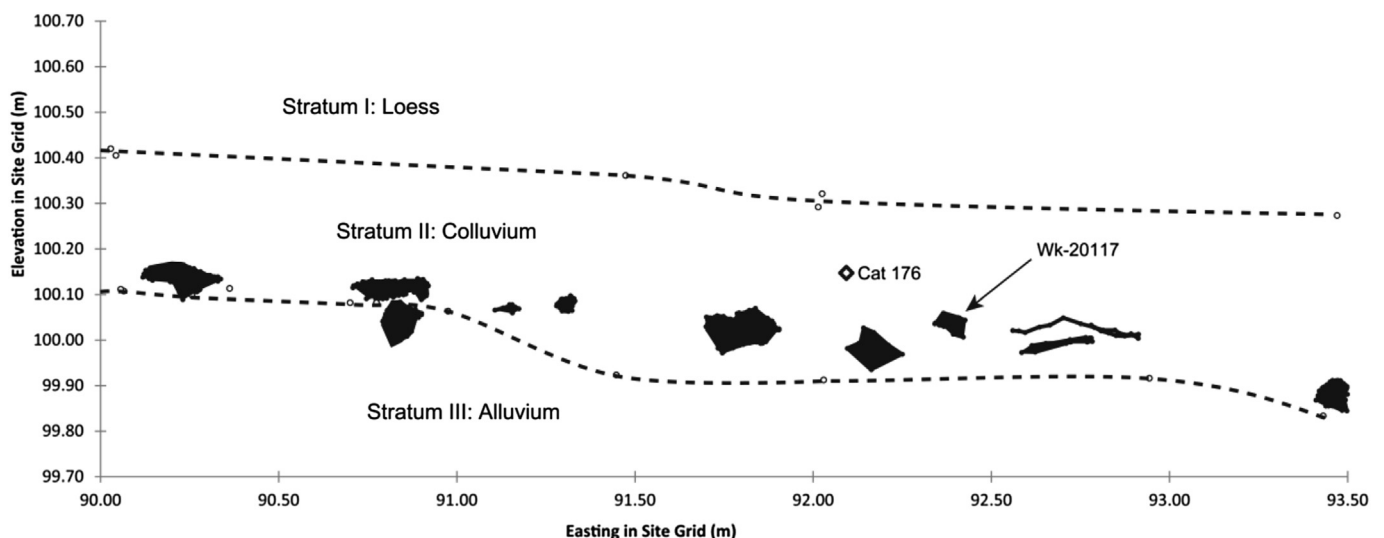
Fig. 2. Elevation backplot (modeled East-West cross-section) showing in situ bones, a possible artifact (catalog no. 176), and strata. Graph shows all total station data for a 30 cm wide strip (500.70−501.00 m North) from 90.0 to 93.5 m East. Black polygons are all bones mapped within these coordinates. Wk-20117 is the location of a bone collagen radiocarbon sample assayed at 13,788 ± 70 RCYBP. Open circles are stratum boundary data points. Both stratum boundaries are wavy (undulating with width greater than depth) and gradual (5−15 cm thick) to diffuse (>15 cm), using National Resource Conservation Service terminology (Schoeneberger et al., 2012).
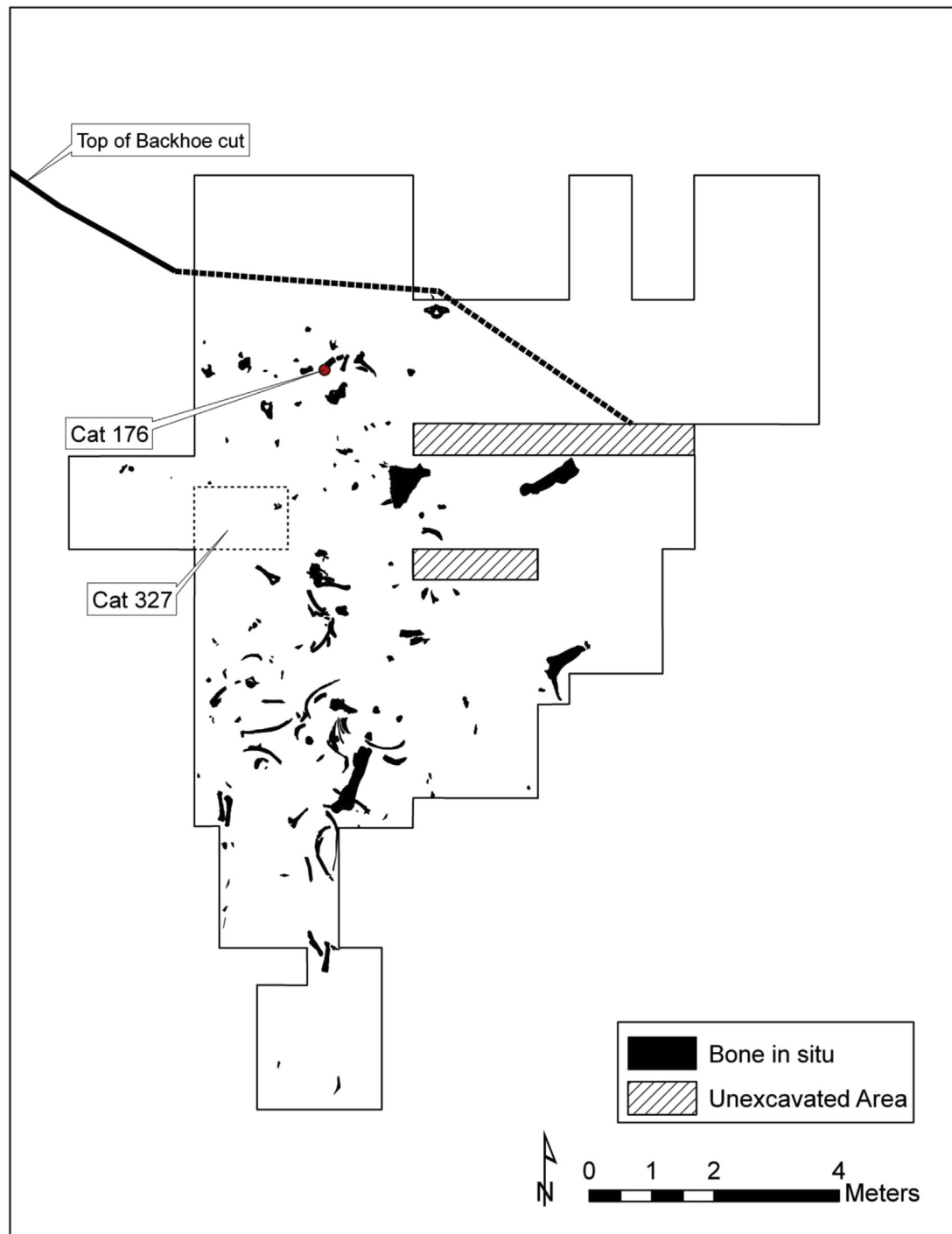
Fig. 3. Excavation block with bone finds and possible artifact locations indicated.

cryptocrystalline silicate (CCS) toolstone sample from the site matrix, and (3) attribute scores from the site toolstone matrix sample and two flintknapped debitage samples. The systematic toolstone matrix sample consisted of all CCS material recovered in the 1/8 inch (0.32 cm) dry screen from 19 excavation units during 2008–2010 field seasons (see Fig. 4). These units constitute 24.9 m³ of excavated sediment, which is about 35% of the total block excavations.

In the excavation screens, the CCS toolstone was chert distinguished fairly readily from the other coarse clasts, principally grainy igneous rocks in shades of gray like andesite, rhyolite, pumice and tuff. The toolstone sample was selected as all rocks with smooth (unpitted) surfaces, no visible grain, and no phenocrysts, including any with luster, non-gray color, or translucency. The sample includes pieces ranging from well-rounded to very angular in shape. The well-rounded CCS is likely underrepresented in the sample because it was more difficult for screeners (field school students) to distinguish in this state than in more angular form.
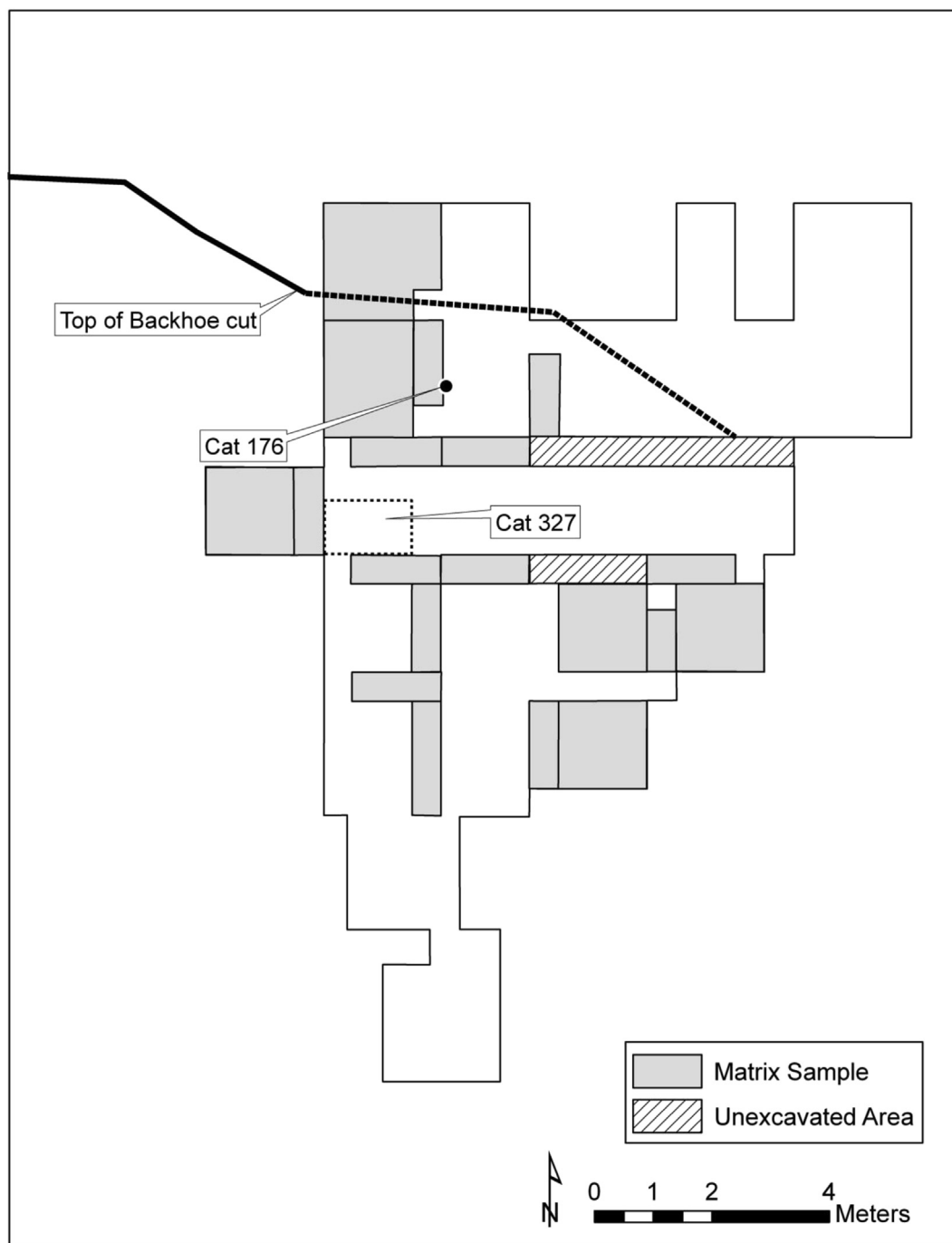
Fig. 4. Excavation block with possible artifact finds and systematic matrix sample units indicated.

The entire CCS toolstone matrix sample ($n = 1167$) was returned to the laboratory, where several attributes were recorded to characterize the nature of toolstone specimens at the site. This was done in two phases, with an initial phase recording general attributes of the entire sample, and a second phase recording technological flaking attributes of the most angular fragments in the sample. The first phase recorded roundness, sphericity, color, translucency, and size grade. This portion of the analysis was completed by graduate student Nate Morse. Roundness was recorded once for the most angular aspect of each specimen as 0.1, 0.3, 0.5, 0.7 or 0.9 (from very angular to well rounded) using the graphic provided by Drevin and Vincent (2002:Fig. 1), which is a subset of the chart provided by

Krumbein (1941). Sphericity was characterized once for the least spherical aspect of each specimen as 0.3, 0.5, 0.7 or 0.9 (from oblong to spherical) using the same graphic provided by Drevin and Vincent (2002:Fig. 1). Color was coded for the dominant material color using the Munsell rock color chart (Munsell Color, 2009). Translucency was listed as opaque (no light penetration) or translucent (any light penetration). Size grade was recorded to approximate sorted screen sizes of 1 inch, ¾ inch, ½ inch, ¼ inch, or 1/8 inch (2.54, 1.91, 1.27, 0.64, and 0.32 cm, respectively) using Andrefsky's (2005:Figure 5.10) size grade chart. For consistency, specimens were turned to the smallest aspect for size grade, and recorded for the largest screen size in which they would be caught.

For specimens recorded as 0.1 (very angular) on roundness, more information was gathered. These specimens were pulled from the complete CCS toolstone matrix sample and verified as 0.1 before further characterization by Terry and Lubinski. The resulting 85 specimens were coded for the lithic debitage attributes noted as helpful in distinguishing geofacts and artifacts (Table 1 attributes A through J, $n = 10$), with the exception of (C) differential weathering of flake scars. The differential weathering attribute was not recorded because it was found to be too subtle for distinction on the small pieces used in this study. In other words, given that most examined specimens were no more than 10 mm in maximum dimension, we found it too difficult to distinguish differences in weathering on flake scars with any confidence. Attribute (J), typically referred to as "dorsal cortex" was adjusted to simply "any cortex" to allow its systematic recording on all specimens, regardless of whether they possessed identifiable dorsal and ventral surfaces. The two possible artifacts (catalog no. 176 and 327 mentioned above) were recorded using the same methods as the systematic toolstone matrix sample. Terry and Lubinski classified specimens jointly, verifying each attribute of each specimen with one another to reduce interobserver error as well as changes in attribute identification during analysis.

The nine attributes recorded for the toolstone matrix sample were used to create a scoring system (Table 2) employed for the two possible artifacts, toolstone matrix sample, and two samples of flintknapped debitage representing cultural material. For dorsal flake scar count, we used three as the cutoff of "few" vs. "multiple." Although it is understood that some authors accept one or two scars as more common in cultural assemblages (Patterson, 1983; Peacock, 1991; Staley, 2006), we prefer the more conservative cutoff of three to account for the findings of Luedtke (1986), whose experiment found 21% of incidentally-produced obsidian flakes exhibited no flake scars, 63% had 1—2 flake scars, and only 16% had 3 or more. Scoring was performed collaboratively by Lubinski and Terry for all samples.

The scoring system is conservative, in that it requires the presence of test-supported features for human-made artifacts from Table 1 to score points, and no attribute is weighted over others. No point is scored for an attribute not meeting the "typical of artifacts" expectation. The potential score for a specimen will vary considerably by debitage type or specimen completeness (e.g., complete flakes, broken flakes, flake fragments, and debris), and also by specimen toolstone character. Thus, complete and broken flakes may score up to 9, while flake fragments, which lack proximal ends and so cannot exhibit the four platform and bulb attributes, may score only 5 at most. Debris, with no identifiable dorsal and ventral surfaces, may score only 1 at most, if it lacks cortex. Complete, fine-grained flakes may score up to 9, while complete, coarse-grained flakes, for which eraillures and fissures are less likely to be observed, may score only 7 at most. Toolstone type and degree of weathering may also affect potential scores, since flintknapping experiments have shown that some raw materials are more likely to fracture and produce specimens with missing proximal ends (e.g., Amick and Mauldin, 1997; Driscoll, 2011). The uneven treatment in the scoring system is intentional. The fact of the matter is that more complete and finer-grained materials are more easily distinguished from geofacts than more fragmentary and coarser-grained materials. The scoring system is conservative by design, making it "harder" for specimens to score as artifacts, and thus provides more confidence that specimens are not geofacts.

The comparative flintknapped sample was derived by collecting all debitage during flintknapping sessions by one of us (Terry) using freehand, hard hammer percussion and two types of CCS raw material. The freehand hard hammer technique was chosen as the most common flintknapping technique (Andrefsky, 2005). The material consisted of one large (6.5 kg) tabular nodule from the

Table 2
Lithic debitage attribute scoring.

Attribute	Typical of geofacts (score 0)	Typical of artifacts (score 1)
Identifiable dorsal & ventral surface	Absent	Present
Platform cortex	NA or Present	Absent
Bulb of percussion	NA or Absent	Present
Bulb of percussion shape	NA or Diffuse	Pronounced
Eraillure scars	NA or Absent	Present
Fissures	NA or Absent	Present
Dorsal flake scar count	NA or <3	3 or more
Dorsal flake scar orientation	NA or Not parallel	Parallel to medial axis
Any cortex	Present	Absent

Ellensburg Formation, obtained from uphill of the study site, and one rounded nodule (0.9 kg) from the Galena Formation of Wisconsin. The Ellensburg Formation nodule was chosen as representative of likely flaking characteristics of the material available nearby, while the Galena Formation nodule was chosen as a different material with a similar texture. For each nodule, cores were reduced until the remaining objective pieces were exhausted, that is, either too small to produce additional flakes, or had no exterior angles that would allow flake removal. Each nodule was reduced by a moderately skilled knapper in a single flintknapping session. All resulting debitage was collected from a 1/8 inch (0.32 cm) screen. This provided an initial sample of 664 pieces from the Galena Formation, and 1543 pieces from the Ellensburg Formation. Because these flintknapped samples were intended for comparison to the 85 very angular matrix samples, given that these initial sample sizes vary 18-fold, and for expediency, we chose to reduce the flintknapped samples for scoring to 100 each. In order to select the flintknapped samples for scoring, we numbered each piece and used a random number generator to pick 100 pieces of each raw material for this analysis. These 200 specimens were scored using Table 2. We then graphically compared our possible artifacts, toolstone matrix sample, and flintknapped debitage sample in the same manner as others who have employed a scoring system (Gillespie et al., 2004; Peacock, 1991; Staley, 2006).

3. Results

3.1. Possible artifacts vs. cultural and geofact attributes

The first possible artifact is catalog no. 176, a flake fragment (Sullivan and Rosen, 1985) which measures 13 mm long by 16 mm wide using the "box" method (Debénath and Dibble, 1994: 19), and 3 mm in maximum thickness. It has all edges broken transversely, except for a feathered lateral left margin and a notched area on the proximal end (Fig. 5). The specimen lacks cortex and has one intact (left) lateral margin parallel with two prominent dorsal arrises, resembling a medial blade fragment. The length of one of these arrises has been removed by a flake scar ~1 mm wide. There are four dorsal flake scars in total. A proximal notch-like feature exhibits 2—3 regular unifacial flake scars ~1 mm in length on the dorsal surface. The specimen is made of red translucent CCS (Munsell color 10R4/6), which under magnification appears as red inclusions in a transparent groundmass (Fig. 6), and is unique among the rocks found in the excavated matrix. No such material was found among the matrix toolstone sample described in Section 3.2 below.

By comparison to the artifact vs. geofact features described above (Table 1), catalog no. 176 exhibits four attributes generally consistent with artifacts: identifiable dorsal and ventral surface, multiple dorsal flake scars, dorsal flake scars parallel to the medial axis, no dorsal cortex. Furthermore, it is made of non-local raw material based on the lack of similar material in the excavated

matrix. The specimen exhibits three attributes generally consistent with geofacts: no bulb of percussion, no eraillures, no fissures, but these are not compelling, being easily accounted for by the lack of a proximal end as expected for flake fragments.

If catalog no. 176 is a geofact, the exotic, parent core stone must have been struck in such a way as to remove the cortex and at least three parallel flakes from its dorsal surface before the object flake was struck parallel to the previous flake removals. The simplest way to account for this would be four closely spaced, subsequent parallel blows to the core stone, which is unexpected for a stone clast whose depositional history involves fluvial and colluvial movement. Additionally, these repeat blows must have occurred without subsequent fluvial/colluvial battering rounding the arrises and forming a noticeable cortex on these surfaces. Alternately, if all of these features could be created from a single blow to the parent core stone during mass wasting, then the geofact hypothesis is more likely, although complicated by its apparent non-local lithology.

Catalog no. 327 is a near-complete specimen classified as a broken flake (Sullivan and Rosen, 1985) due to breakage of a portion of the distal margin. The specimen measures 9 mm long by 12 mm wide using the "box" method, and is 2 mm in maximum thickness. It exhibits a flat platform with an angle of 55°, an eraillure scar on a diffuse bulb of percussion, 5 dorsal flake scars (including two 1-mm-long scars adjacent to the platform), feather termination, and no cortex (Fig. 7). The specimen is made of translucent tan CCS (Munsell color 10YR8/6) visually similar to a small proportion of the material found in the excavation matrix. (Of the matrix toolstone sample described in Section 3.2 below, 29 of 1167 specimens (2.5%) were translucent and 5YR or 10YR, while 7 (0.06%) were translucent and 10YR8/6.)

By comparison to the artifact vs. geofact features described above (Table 1), catalog no. 327 exhibits six attributes generally consistent with artifacts: identifiable dorsal and ventral surface, platform cortex absent, bulb of percussion, eraillure scar, multiple dorsal flake scars, and no dorsal cortex. The specimen exhibits three attributes generally consistent with geofacts: diffuse bulb of percussion, no fissures, no flake scars parallel to medial axis. Furthermore, it is made of apparently local raw material, based on the presence of similar material in the excavated matrix.
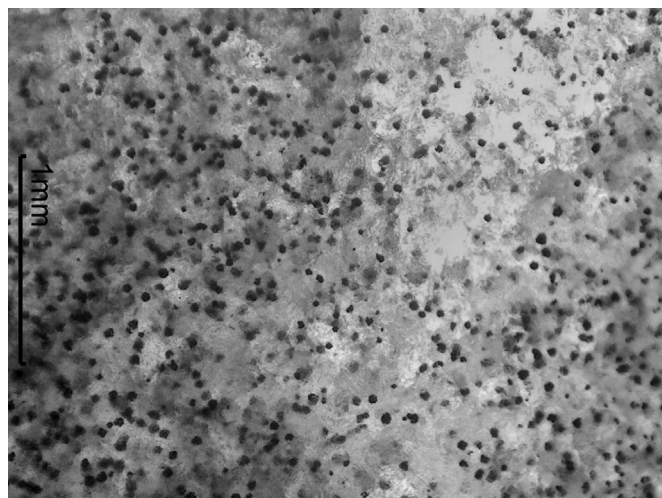

Fig. 6. Photomicrograph of raw material for catalog no. 176.

If catalog no. 327 is a geofact, the parent core stone must have been struck in such a way as to remove the cortex and at least three non-parallel flakes from its dorsal surface before the object flake was struck. The simplest way to account for this would be four subsequent blows to the core stone at somewhat random orientations, which is entirely possible for a stone clast whose depositional history involves fluvial and colluvial movement. However, a depositional setting with sufficient clasts and energy to remove four flakes would be expected to also result in additional fluvial/colluvial battering that did not remove flakes and could be recognized on the specimen. Catalog no. 327 did not exhibit such battering. Alternately, if all of these features could be created from a single blow to the parent core stone, then the geofact hypothesis is more likely.

These two possible artifacts have both artifact and geofact attributes, although the expressed artifact attributes are more frequent and compelling, plus the geofact attributes are not diagnostic of natural pieces. This first comparison does not provide compelling support for the geofact hypothesis. However, a stronger
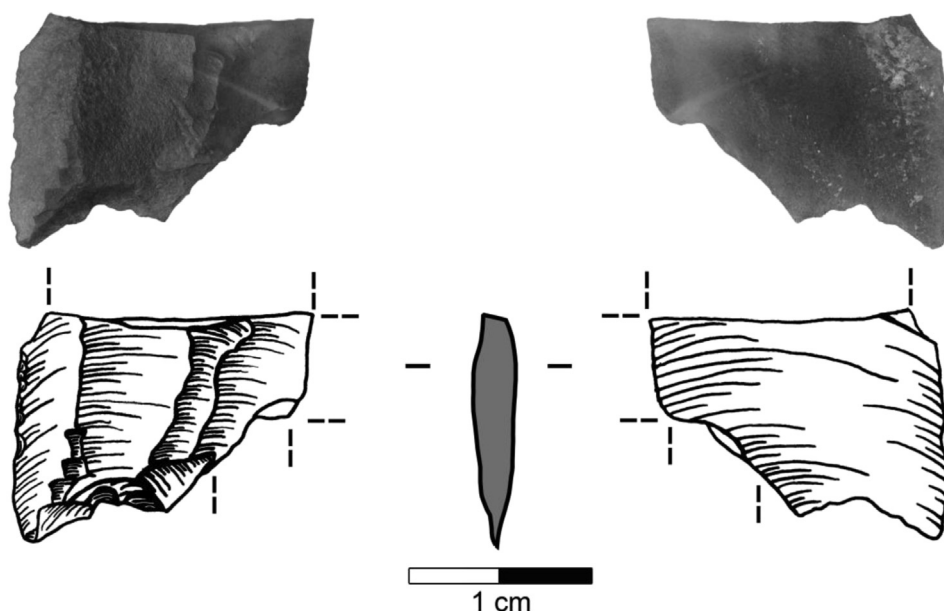

Fig. 5. Possible artifact (catalog no. 176). Distal end is to top and dorsal and ventral surfaces are on left and right sides, respectively. Fractured margins are indicated between pairs of parallel dashed lines.

evaluation requires consideration of the natural toolstone present in the site matrix.

3.2. Possible artifacts vs. matrix toolstone

The systematic toolstone matrix sample consisted of 1167 specimens. When divided by excavation volume, this is a density of 46.9 pieces/m^3. The majority were moderately rounded (peak at 0.3), moderately spherical (peak at 0.5), yellow-red, opaque, and from the 1/8 inch (0.32 cm) or 1/4 inch (0.64 cm) fraction (see Fig 8, Table 3). The specimens ranged up to 56 mm in maximum dimension. Given the small size of most specimens (99% from <1 inch [2.54 cm] fraction), the quality and knappability of the site matrix toolstone is uncertain. Only 7% of the sample ($n = 85$) was very angular (roundness 0.1).

These 85 very angular specimens were examined further, summarized in Table 4. This angular subsample was primarily debris with no bulb of percussion, no fissures, no eraillures, and no dorsal flake scars. About half of them ($n = 38$) exhibited a hard pebble cortex, while the remainder retained no visible cortex. Two of these specimens exhibited identifiable dorsal and ventral surfaces and would be classified as a flake fragment and complete flake under the criteria of Sullivan and Rosen (1985), although the majority of features on both specimens are more consistent with geofacts than artifacts. Additionally, both specimens are composed of opaque CCS with a yellow-red hue, a typical color combination which made up 70% of the toolstone matrix sample. The first specimen, recorded as a flake fragment, exhibited identifiable dorsal and ventral surfaces, but no other cultural feature listed in Table 1. Because it lacks a proximal end, it did not exhibit cortex on the platform, bulb of percussion, pronounced bulb shape, or eraillures. The specimen also lacked fissures, multiple flake scars, or flake scars parallel to the medial axis. The second specimen, recorded as a complete flake, exhibited a bulb of percussion and multiple dorsal flake scars typical of artifacts, but no other typical cultural feature. This specimen also exhibited cortex on the platform, a diffuse bulb of percussion, no eraillures, no fissures, and dorsal flake scars not parallel to medial axis, all typical of geofacts. Although both of these matrix sample specimens exhibited identifiable dorsal and ventral surfaces, they would be discarded as

geofacts by many researchers due to their paucity of purportedly cultural features and the fact they appear composed of a commonly-occurring raw material in the site matrix. We argue that neither of these two specimens nor any of the remaining matrix sample are cultural artifacts.

When compared to this systematic toolstone matrix sample, the two possible artifacts are unusual, having many attributes with very low frequencies in the sample. This is apparent for both toolstone attributes (angularity, translucency, color, size) and fracture morphology attributes (identifiable dorsal and ventral surface, multiple dorsal flake scars, eraillure scars, dorsal flake scar orientation). Even considering only the toolstone attributes of material roundness (0.1) and translucency (translucent) like the two possible artifacts, the matrix sample has only 2 examples (0.2% of the total). In terms of only color and translucency, there are some (29; 2.5%) yellow-red hue translucent pieces like catalog no. 327, but very few (3; 0.3%) red hue translucent pieces in the matrix sample like catalog no. 176. In fact, both possible artifacts are unique in raw material if considering size grade, translucency, and roundness. Catalog no. 327 is translucent like 12 examples (1.0%) of 1/4 inch (0.64 cm) size grade pieces, but none of these matrix examples are very angular (roundness 0.1) like catalog no. 327. Catalog no. 176 is of a larger size grade (3/4 inch or 1.91 cm) than any matrix example of a translucent piece of any roundness. Additionally, as noted above in Section 3.1, this possible artifact is distinct microscopically compared to the matrix sample, being the only one composed of a set of red inclusions in a clear groundmass.

In addition to the unusual toolstone, the possible artifacts are unusual in terms of their fracture morphology. Both possible artifacts have easily identifiable dorsal and ventral surfaces, but only 2 of the 1167 pieces (0.2%) in the toolstone matrix sample share this trait. Both possible artifacts also have multiple flake scars on the dorsal side, a trait shared with only one of the matrix sample pieces (0.1%). In fact, the possible artifacts are unique compared to the matrix sample in terms of at least one flake attribute. Catalog no. 176 has multiple flake scars parallel to the long axis, unlike any of the matrix sample. Catalog no. 327 has an eraillure scar, and can be identified as a broken flake, all features not found in the matrix sample.
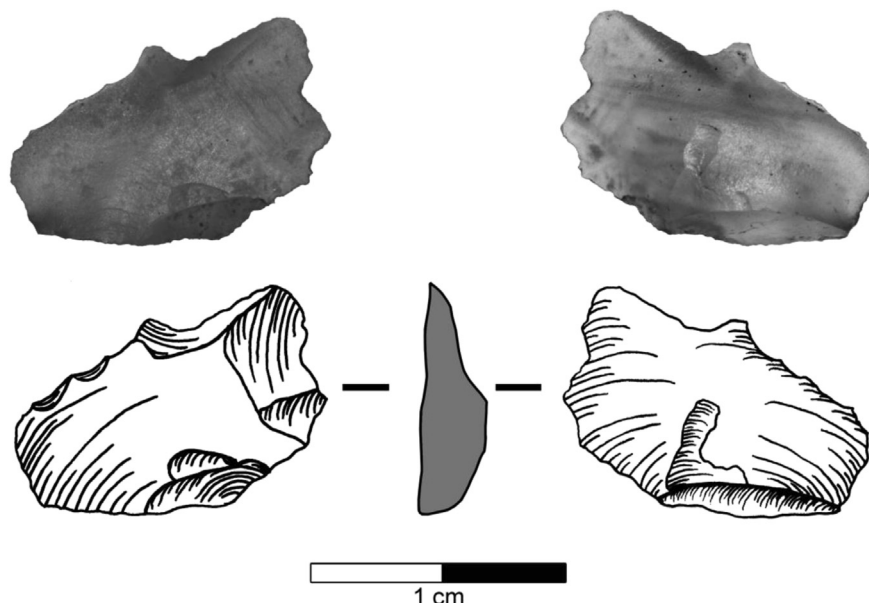

Fig. 7. Possible artifact (catalog no. 327). Distal end is to top and dorsal and ventral surfaces are on left and right sides, respectively.

Fig. 8. An example of the matrix toolstone samples. This one has the most specimens (catalog no. 713, *n* = 57).

In sum, the possible artifacts have features that are rare or absent in the matrix sample. This second comparison does not provide compelling support for the geofact hypothesis. However, a stronger evaluation combines consideration of attribute features and comparison to the matrix sample using a scoring approach.

3.3. Possible artifacts vs. attribute scores

As a third way to evaluate the possible artifacts, we compare debitage attribute scores for the two possible artifacts against the

very angular (roundness 0.1) toolstone matrix sample and the comparative flintknapped debitage samples. Table 5 provides a summary of the two flintknapped samples and the matrix toolstone sample, as well as scores for the two possible artifacts. Attributes for all three samples were converted to scores using Table 2, and the resulting distribution is shown in Fig. 9. In the flintknapped debitage samples, only two of the characteristic attributes of cultural debitage occurred in high proportions, namely identifiable dorsal and ventral surface (76–93%) and lack of any cortex (74–93%). Three attributes were common (≥36%): no platform cortex, bulb of percussion, three or more dorsal flake scars. The remaining four attributes were rare (≤26%): pronounced bulb of percussion, eraillures, fissures, parallel dorsal flake scars.

Fig. 9 shows the distribution of the toolstone matrix samples and the scores of the two possible artifacts. The matrix sample distribution is strongly unimodal, with a peak at score 1 and a maximum at score 3. The flintknapped debitage samples are quite different, with peaks at score 1 or 2, similar proportions through score 5, and maximums at score 8. While both the matrix and flintknapped debitage sample score distributions have similar peaks, they are quite different in their medians (1 vs. 3 and 3.5) and ranges (3 vs. 7 and 8). The two possible artifacts (catalog nos. 176 and 327) have scores of 4 and 6 respectively. Both of these scores are well within the flintknapped sample distributions and are strongly separated from the matrix sample distribution. These data do not support the geofact hypothesis.

4. Discussion and conclusions

This study attempted to distinguish geofacts from cultural debitage at the Wenas Creek Mammoth site. In doing so, we have generated some experimental data that provide useful generalizations. For instance, most attributes were not common enough to provide compelling support for the geofact vs. artifact distinction by themselves (Table 5). For example, eraillures and fissures occurred in only 2–14% of flintknapped debitage specimens, proportions too rare to be useful in isolation. Of the attributes we recorded, only the presence of identifiable dorsal and ventral surfaces was common in the flintknapped debitage sample (76–93%), while very rare in the angular matrix sample (2%). However, even this attribute was not diagnostic and could not be used alone, since two of the 85 angular matrix samples exhibited this trait. For this reason, like other researchers (e.g., Bradbury, 2001; Patterson,

Table 3
Matrix toolstone sample summary (*n* = 1167).

Variable	Values (%)				
	0.1	0.3	0.5	0.7	0.9
Roundness[a]	7	48	33	11	1
	0.1	0.3	0.5	0.7	0.9
Sphericity[b]	NA	32	36	28	4
	Yellow	YR	Red	Other	
Color	14	72	14	<1	
	Opaque	Translucent			
Translucency	97	3			
	1/8 in	¼ in	½ in	¾ in	1 in
Size Grade	50	42	6	1	1

[a] from 0.1 (very angular) to 0.9 (well-rounded)
[b] from 0.3 (oblong) to 0.9 (spherical)

Table 4
Matrix toolstone very angular sample summary (*n* = 85).

Variable	Values (Count)			
	Debris	Flake Frag.	Broken Fl.	Complete Fl.
Debitage class	83	1	0	1
	Yes	No	NA	
Platform cortex	1	0	84	
	Yes	No (NA)		
Bulb	1	84		
	Diffuse	Pronounced	NA	
Bulb shape	1	0	84	
	Yes	No	NA	
Eraillures	0	1	83	
	Yes	No		
Fissures	0	84		
	>1 (n=6)	1	0 (NA)	
Dorsal scar count	1	1	83	
	Parallel	Not parallel	NA	
Dorsal scar orientation	0	2	83	
	Yes	No		
Any cortex	38	47		

1983; Peacock, 1991; Staley, 2006), we recommend the use of a suite of attributes, especially in a scoring system like used here.

In the particular case of the Wenas Creek Mammoth site, the two possible artifacts were found more consistent with artifacts than geofacts in three ways. First, the possible artifacts share more attributes typically associated with artifacts than geofacts. Second, they are quite different from the natural toolstone fragments found in the site matrix, in terms of both raw material and fracture morphology. Third, they are more similar to flintknapped debitage samples than to the site toolstone matrix sample based on scored technological attributes. Thus, in a probabilistic sense, they are more likely to be artifacts than geofacts, being much more similar to culturally-produced materials than to the naturally-occurring site matrix sample. Of course this provisionally-accepted hypothesis is far from exhaustive, and other possibilities remain, such as these are in fact geofacts with different characteristics than the matrix sample. An alternative hypothesis is that they are zoofacts, perhaps like the chert pseudo-flakes created in elephant and bison trampling experiments reported by Lopinot and Ray (2007). Although we did not undertake a detailed evaluation of this hypothesis, all of the 230+ flake-like specimens created in those experiments exhibited dorsal cortex (Lopinot and Ray, 2007: 776, 777), unlike the specimens here. That work did not report any of the other attributes in our Table 1.

Table 5
Attribute summary for flintknapped samples, matrix samples, and possible artifacts.

Variable	Samples			Cat. 176 Scores	Cat. 327 Scores
	Galena (%)	Ellensburg (%)	Matrix (%)		
Entire sample (sample size)	N = 100	N = 100	N = 85		
Identifiable dorsal surface (yes)	93	76	2	1	1
Platform cortex (none)	42	36	0	0	1
Bulb (present)	51	49	1	0	1
Bulb shape (pronounced)	9	21	0	0	0
Eraillures (present)	2	7	0	0	1
Fissures (present)	10	14	0	0	0
Dorsal scar count (3+)	44	47	1	1	1
Dorsal scar orientation (parallel to long axis)	26	21	0	1	0
Any cortex (none)	74	93	47	1	1
When bulb is present (sample size)	N = 51	N = 49	N = 1	N = 0	N = 1
Platform cortex (none)	82	69	0		
Bulb shape (pronounced)	18	43	0		
Eraillures (present)	4	12	0		
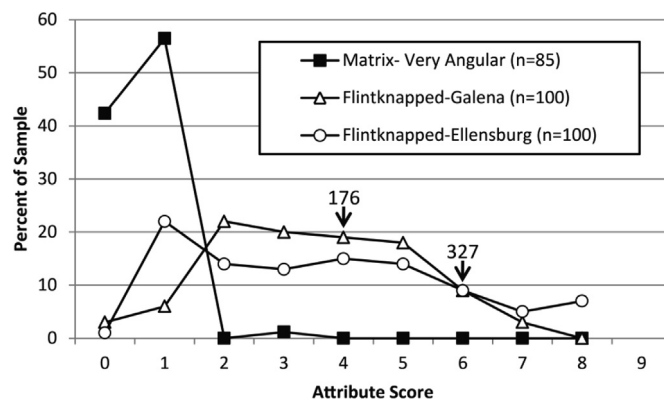
Fig. 9. Attribute scores of the toolstone matrix and experimental debitage samples. Attribute scores for the two possible artifacts (catalog nos. 176 and 327) are indicated.

Some claims for artifacts at other sites have faltered on the basis that there was a large amount of site matrix raw material specimens from which to selectively choose the "artifacts" (e.g., Duvall and Venner, 1979; Haynes, 1973; Roebroeks and van Kolfschoten, 1994). The Wenas Creek Mammoth site also has toolstone in the site matrix, but the purported artifacts are quite different from the matrix toolstone specimens. Additionally, the presence of toolstone at a site and the presence of artifacts at a site may be independent phenomena, and should not be assumed causally linked. After all, mixtures are common at other locales with naturally-occurring toolstone, including quarry sites (e.g., Andrews et al., 2004).

The geological context of the Wenas Creek site does not support a geofact interpretation. That is, the depositional history of the sediment bearing the purported artifacts does not appear to reflect the energy necessary for natural flake production, at least not for specimens with multiple dorsal flake scars lacking any cortex. This claim is supported by the low incidence of specimens with identifiable dorsal and ventral surfaces in the systematic matrix sample, and the low incidence of coarse clasts in the stratum. Additionally, the bones from the same stratum lack the edge rounding and polishing expected for high energy fluvial environments (see Fernández-Jalvo and Andrews, 2003). This latter point is important, given the fragility of bone cortex surfaces compared to cryptocrystalline silicates, unless the bones were deposited first in a fine-grained colluvium, and the possible artifacts later in coarser colluvium that could not be stratigraphically separated in the field.

The hypothesis that the Wenas Creek specimens are geofacts is not well supported by the testing protocol used here. However, since other alternative hypotheses have not been systematically tested, it may be premature to accept them as artifacts. Instead, they are best thought of as possible or probable artifacts. Other terms have been proposed for such ambiguous specimens, such as *equifacts* (West, 1983) or *incerto-facts* (Roebroeks and Stapert, 1986). As implied in the term *equifact*, the concept of equifinality may be pertinent here. As applied by many archaeologists today, equifinality means that more than one causal process may result in a particular outcome, so that we may be unable to distinguish between several possible causes of an archaeological pattern (see Lyman, 2004). Some may argue that the process that created the Wenas Creek Mammoth site specimens cannot ultimately be distinguished between our two causal alternatives with certainty, such pieces always being the victim of equifinality. However, we prefer to emphasize that our test has provided a useful if partial solution to the problem and encourage further attempts at finding pertinent analytical tools, as Andrefsky (2013) urges, rather than "assume a fatalistic attitude and not develop better methods" (Lyman, 2004: 22).

Given the stratigraphy and dating of the site, reasonable interpretations can range from the idea that there are two possible artifacts from 17 ka to the idea that there are two later, genuine artifacts intruded into a 17 ka paleontological site. If these two specimens are in fact human-made artifacts, and if they are associated with the well-dated 17 ka mammoth and bison bones, given the relative dearth of pre-Clovis age sites, it would have implications for our understanding of the peopling of the Americas. However, the evidence from the Wenas Creek Mammoth site is not sufficiently strong to make this assertion, given the uncertainty about whether the specimens are in fact artifacts and the uncertainty about their age. The site and its lithic specimens retain an ambiguous position in our understanding of prehistory. While this finding may be unsatisfying, we believe it represents a well-supported and intellectually honest evaluation of the site and its lithic specimens. Additional work in progress or planned for the site includes microwear analysis of the possible lithic artifacts,

taphonomic analysis of the bones for possible cutmarks, and microstratigraphic investigation of sediments.

We believe we have shown a method of testing for geofacts that may be more widely used by others. Other researchers could use the scoring system given in Table 2 on their own uncertain debitage/geofact samples and additional flintknapped samples, and compare scores to our flintknapped and matrix toolstone samples. For more rigorous testing, comparison of scores to site-specific matrix toolstone samples would be required. All of these possible uses require data on the nine debitage attributes given in Table 2. Since we were unable to find these data for other purportedly early North American sites, such an evaluation remains in the future.

Naturally, the method used here could be improved with additional work. Its replicability could be examined to determine if the selected debitage attributes can be recorded in a comparable way by multiple investigators (cf. Beck and Jones, 1989; Potts, 2012). Additional experimental or actualistic data could also be added to Table 2 to provide a more secure basis for geofact vs. artifact distinctions, perhaps adding or removing attributes in light of these data.

The method described here may be better for showing that questionable specimens are likely geofacts than artifacts, since most researchers would conservatively dismiss specimens more similar to geofacts as geofacts, but interpret specimens more similar to artifacts as artifacts only provisionally. Regardless, we hope that this proposed approach will provide another tool for interpreting small samples of possible debitage, and thus assist in the systematic and objective evaluation of such finds.

References

Ackerman, S., 1989. European prehistory gets even older. Science 246, 28–30.
Amick, D.S., Mauldin, R.P., 1997. Effects of raw material on flake breakage patterns. Lithic Technol. 22, 18–32.
Andrefsky Jr., W., 2005. Lithics: Macroscopic Approaches to Analysis, second ed. Cambridge University Press, Cambridge.
Andrefsky Jr., W., 2013. Fingerprinting flake production and damage processes: toward identifying human artifact characteristics. In: Graf, K.E., Ketron, C.V., Water, M.R. (Eds.), Paleoamerican Odyssey. Center for the Study of the First Americans, College Station, Texas, pp. 415–428.
Andrews, B.W., Murtha Jr., T.M., Scheetz, B., 2004. Approaching the Hatch Jasper Quarry from a technological perspective: a study of prehistoric stone tool production in central Pennsylvania. MidCont J. Archaeol. 29, 63–101.
Ascher, R., Ascher, M., 1965. Recognizing the emergence of man. Science 147, 243–250.
Barnes, A.S., 1939. The differences between natural and human flaking on prehistoric flint implements. Am. Anthropol. 41, 99–112.
Beck, C., Jones, G.T., 1989. Bias and archaeological classification. Am. Antiq. 54, 244–262.
Bentley, R.D., Campbell, N.P., 1983. Geologic Map of the Yakima Quadrangle, Washington (1:62,500). Washington State Department of Natural Resources, Olympia.

Bradbury, A.P., 2001. Modern or prehistoric: experiments in distinguishing between culturally and mechanically produced chipped stone artifacts. N. Am. Archaeol. 22, 231–258.

Buehler, K., 2003. Human and naturally modified chipped stone items from the Burnham site. In: Wyckoff, D.G., Theler, J.L., Carter, B.J. (Eds.), The Burnham Site in Northwestern Oklahoma: Glimpses beyond Clovis? Sam Noble Oklahoma Museum of Natural History, Norman, Oklahoma, pp. 207–231.

Byers, D.A., 2005. Taphonomic analysis, associational integrity, and depositional history of the Fetterman mammoth, eastern Wyoming, U.S.A. Geoarchaeology 17, 417–440.

Collins, M.B., 1997. The lithics from Monte Verde, a descriptive-morphological analysis. In: Dillehay, T.D. (Ed.), Monte Verde: a Late Pleistocene Settlement in Chile, The Archaeolgocial Context and Interpretation, vol. 2. Smithsonian Institution Press, Washington, D.C, pp. 383–506.

Cotterell, B., Kamminga, J., 1987. The formation of flakes. Am. Antiq. 52, 675–708.

Debénath, A., Dibble, H.L., 1994. Handbook of Paleolithic Typology Volume One: Lower and Middle Paleolithic of Europe. University of Pennsylvania Museum, Philadelphia.

de Mortillet, G., 1883. Le Préhistorique Antiquité de L'homme. C. Reinwald, Paris. Available online at: http://www.archive.org/details/leprhistorique00mort (accessed 10.05.11.).

Dincauze, D.F., 1984. An archaeo-logical evaluation of the case for pre-Clovis oc-cupations. In: Wendorf, F., Close, A.E. (Eds.), Advances in World Archaeology, vol. 3. Academic Press, Orlando, Florida, pp. 275–323.

Drevin, G.R., Vincent, L., 2002. Granulometric determination of sedimentary rock particle roundness. In: Talbot, H., Beare, R. (Eds.), Mathematical Morphology: Proceedings of the VIth International Symposium ISMM 2002. CSIRO Publish-ing, Collingwood, Australia, pp. 315–325.

Driscoll, K., 2011. Vein quartz in lithic traditions: an analysis based on experimental archaeology. J. Archaeol. Sci. 38, 734–745.

Dunbar, J.S., 2006. Paleoindian archaeology. In: Webb, S.D. (Ed.), First Floridians and Last Mastodons: the Page-Ladson Site in the Aucilla River. Springer, Dordrecht, The Netherlands, pp. 403–435.

Duvall, J.G., Venner, W.T., 1979. A statistical analysis of the lithics from the Calico site (SBCM 1500A), California. J. Field Archaeol. 6, 455–462.

Ellen, R., Muthana, A., 2013. An experimental approach to understanding the "eolithic" problem: cultural cognition and the perception of plausibly anthropic artifacts. Lithic Technol. 38, 109–123.

Eren, M.I., Boehm, A.R., Morgan, B.M., Anderson, R., Andrews, B., 2011. Flaked stone taphonomy: a controlled experimental study of the effects of sediment consolidation on flake edge morphology. J. Taphon. 9, 201–217.

Fernández-Jalvo, Y., Andrews, P., 2003. Experimental effects of water abrasion on bone fragments. J. Taphon. 1, 147–163.

Fosha, M., Holen, S., Donohue, J., 2012. Virgil Schulz mammoth. Newsl. S. D. Archaeol. Soc. 42 (2), 1–4.

Frison, G.C., Todd, L.C., 1986. The Colby Mammoth Site: Taphonomy and Archae-ology of a Clovis Kill in Northern Wyoming. University of New Mexico Press, Albuquerque.

Fuld, K.A., Ozbun, T.L., Jenkins, S., 2014. Pleistocene Megafauna at Williams North-west Pipeline's Anomaly 20940, Linn County, Oregon. Archaeological In-vestigations Northwest. Report No. 3167.

Gillespie, J.D., Tupakka, S., Cluney, C., 2004. Distinguishing between naturally and culturally flaked cobbles: a test case from Alberta, Canada. Geoarchaeology 19, 615–633.

Grayson, D.K., 1986. Eoliths, archaeological ambiguity, and the generation of "middle range" research. In: Meltzer, D.J., Fowler, D.W., Sabloff, J.A. (Eds.), American Archaeology Past and Future: a Celebration of the Society for American Archae-ology 1935–1985. Smithsonian Institution Press, Washington, D.C, pp. 77–133.

Hamilton, T.M., 1996. The Miami mastodon, 23SA212. Mo. Archaeol. 54, 79–88.

Haynes, C.V., 1969. The earliest Americans. Science 166, 709–715.

Haynes, C.V., 1973. The Calico site: artifacts or geofacts? Science 181, 305–310.

Hosfield, R.T., Chambers, J.C., 2003. Flake modifications during fluvial trans-portation: three cautionary tales. Lithics 24, 57–65.

Joyce, D.J., 2013. Pre-Clovis megafauna butchery sites in the western Great Lakes region, USA. In: Graf, K.E., Ketron, C.V., Water, M.R. (Eds.), Paleoamerican Od-yssey. Center for the Study of the First Americans, College Station, Texas, pp. 467–483.

Krumbein, W.C., 1941. Measurement and geological significance of shape and roundness of sedimentary particles. J. Sediment. Petrol. 11, 64–72.

Leakey, L.S.B., Simpson, R.D., Clements, T., 1968. Archaeological excavations in the Calico Mountains, California: preliminary report. Science 160, 1022–1023.

Lillquist, K., Lundblad, S., Barton, B.R., 2005. The Moxee City (Washington) mammoth: morphostratigraphic, taphonomic, and taxonomic considerations. West. N. Am. Nat. 65, 417–428.

Lopinot, N.H., Ray, J.H., 2007. Trampling experiments in the search for the earliest Americans. Am. Antiq. 72, 771–782.

Lubinski, P.M., Barton, B.R., Lillquist, K., Uebelacker, M., Shapley, J.T., 2007. The Late-Glacial Wenas Creek mammoth site (45YA1083) in central Washington. Curr. Res. Pleistocene 24, 178–180.

Lubinski, P.M., Feathers, J., Lillquist, K., 2014. Single-grain luminescence dating of sediment surrounding a possible Late Pleistocene artifact from the Wenas Creek mammoth site, Pacific Northwest, USA. Geoarchaeology 29, 16–32.

Lubinski, P.M., McCutcheon, P.T., Lillquist, K., Uebelacker, M., Barton, B.R., Shapley, J.T., 2009. Possible lithic artifacts from 2005–2007 excavations at the Wenas Creek mammoth site. Curr. Res. Pleistocene 26, 85–86.

Luedtke, B.E., 1986. An experiment in natural fracture. Lithic Technol. 15, 55–60.

Lyman, R.L., 2004. The concept of equifinality in taphonomy. J. Taphon. 2, 15–26.

Meltzer, D.J., 2004. Peopling of North America. In: Gillespie, A.R., Porter, S.C., Atwater, B.F. (Eds.), The Quaternary Period in the United States. Elsevier, Amsterdam, pp. 539–563.

Meltzer, D.J., Adovasio, J.M., Dillehay, T.D., 1994. On a Pleistocene human occupation at Pedra Furada, Brazil. Antiquity 68, 695–714.

Miller, F., Powell, J., 1997. An Archaeologist's View of Eastern Washington Through the Eyes of a Geologist. Paper Presented at the Northwest Anthropol. Confer-ence, Ellensburg, Washington.

Munsell Color, 2009. Geological Rock-Color Chart. [PDF version of 115-chip chart.] Electronic Document. www.munsellstore.com/files/CIPA00011%5C599.pdf (accessed February, 2010).

Munson, P.J., Frye, J.C., 1965. Artifact from deposits of mid-Wisconsin age in Illinois. Science 150, 1722–1723.

Nash, D.T., 1993. Distinguishing stone artifacts from naturefacts created by rockfall processes. In: Goldberg, P., Nash, D.T., Petraglia, M.D. (Eds.), Formation Pro-cesses in Archaeological Context. Prehistory Press, Madison, Wisconsin, pp. 125–138.

Oakley, K.P., 1967. Man the Toolmaker. Trustees of the British Museum, London.

O'Connell, J.F., Allen, J., 2004. Dating the colonization of Sahul (Pleistocene Australia-New Guinea): a review of recent research. J. Archaeol. Sci. 31, 835–853.

Patterson, L.W., 1983. Criteria for determining the attributes of man-made lithics. J. Field Archaeol. 10, 297–307.

Payen, L.A., 1982. Artifacts or geofacts at Calico: application of the Barnes test. In: Ericson, J.E., Taylor, R.E., Berger, R. (Eds.), Peopling of the New World. Ballena, Los Altos, California, pp. 193–201.

Peacock, E., 1991. Distinguishing between artifacts and geofacts: a test case from eastern England. J. Field Archaeol. 18, 345–361.

Potts, T.L., 2012. Low-quality quartz and implications for technological inferences. In: Carr, P.J., Bradbury, A.P., Price, S.E. (Eds.), Contemporary Lithic Analysis in the Southeast: Problems, Solutions, and Interpretations. University of Alabama Press, Tuscaloosa, pp. 111–127.

Reeves, B.O.K., 1980. Fractured cherts from Pleistocene fossiliferous beds at Medi-cine Hat, Alberta. In: Browman, D.L. (Ed.), Early Native Americans: Prehistoric Demography, Economy, and Technology. Mouton, New York, pp. 83–98.

Richards, R.L., Whitehead, D.R., Cochran, D.R., 1987. The Dollens mastodon (*Mammut americanum*) locality, Madison County, east central Indiana. Proc. Indiana Acad. Sci. 97, 571–581.

Roebroeks, W., Stapert, D., 1986. On the "Lower Paleolithic" site La Belle Roche: an alternative interpretation. Curr. Anthr. 27, 369–371.

Roebroeks, W., van Kolfschoten, T., 1994. The earliest occupation of Europe: a short chronology. Antiquity 68, 489–503.

Sanchez, G., Holliday, V.T., Gaines, E.P., Arroyo-Cabrales, J., Martínez-Tagüena, N., Kowler, A., Lange, T., Hodgins, G.W., Mentzer, S.M., Sanchez-Morales, I., 2014. Human (Clovis)–gomphothere (*Cuvieronius* sp.) aassociation ~13,390 calibrated yBP in Sonora, Mexico. Proc. Natl. Acad. Sci. 111, 10972–10977.

Saunders, J., Daeschler, E.B., 1994. Descriptive analyses and taphonomical obser-vations of culturally-modified mammoths excavated at "the gravel pit," near Clovis, New Mexico in 1936. Proc. Acad. Nat. Sci. Phila 145, 1–28.

Schoeneberger, P.J., Wysocki, D.A., Benham, E.C., Soil Survey Staff, 2012. Field Book for Describing and Sampling Soils (Version 3.0). National Soil Survey Center, Lincoln, Nebraska.

Sellards, E.H., 1938. Artifacts associated with fossil elephant. Geol. Soc. Am. Bull. 49, 999–1010.

Shea, J.J., 2010. Stone Age visiting cards revisited: a strategic perspective on the lithic technology of early Hominin dispersal. In: Fleagle, J.G., Shea, J.J., Grine, F.E., Baden, A.L., Leakey, R.E. (Eds.), Out of Africa I: the First Hominin Colonization of Eurasia. Springer, New York, pp. 47–64.

Smith, G.A., 1988. Sedimentology of proximal to distal volcaniclastics dispersed across an active foldbelt: Ellensburg Formation (late Miocene), central Wash-ington. Sedimentology 35, 953–977.

Staley, D.P., 2006. Shadow of a doubt or doubtful shadows: small-scale low-density lithic scatters and agrifacts. N. Am. Archaeol. 27, 175–199.

Stuiver, M., Reimer, P.J., Reimer, R., 2012. CALIB Radiocarbon Calibration, pc Version 6.1.1. Available at: http://calib.qub.ac.uk/calib/ (accessed 04.07.12.).

Sullivan III, A.P., Rosen, K.C., 1985. Debitage analysis and archaeological interpre-tation. Am. Antiq. 50, 755–779.

Waitt Jr., R.B., 1980. About forty Last-Glacial Lake Missoula jökulhlaups through southern Washington. J. Geol. 88, 653–679.

Waitt Jr., R.B., 1985. Case for periodic, colossal jökulhlaups from Pleistocene glacial Lake Missoula. Geol. Soc. Am. Bull. 96, 1271–1286.

Warren, S.H., 1905. On the origin of eoliths. Man (Lond) 5, 179–183.

West, F.H., 1983. The antiquity of man in America. In: Porter, S.C. (Ed.), Late-Qua-ternary Environments of the United States, The Late Pleistocene, vol. 1. Uni-versity of Minnesota Press, Minneapolis, pp. 364–382.

Wilson, M.C., Burns, J.A., 1999. Searching for the earliest Canadians: wide corridors, narrow doorways, small windows. In: Bonnishsen, R.A., Turnmire, K.L. (Eds.), Ice Age People of North America: Environments, Origins and Adaptations. Oregon State University Press, Corvallis, pp. 213–248.

Wyckoff, D.G., Theler, J.L., Carter, B.J. (Eds.), 2003. The Burnham Site in North-western Oklahoma: Glimpses Beyond Clovis?. Sam Noble Oklahoma Museum of Natural History, Norman, Oklahoma.